GENERAL NOTES.

Some Common Inaccuracies in the Application of the Method of Least Squares.—It is gratifying to notice how wide-spread has become the use of the method of least squares during recent years. When the method was new it found favor only in astronomy and geodesy. Nowadays it is not only used in all kinds of physical problems, but one is no longer surprised to see it applied to problems in biology, in psychology, and even in sociology. Perhaps because of this wider field, or perhaps because of the pitfalls inherent to any branch of mathematical probability, it is certainly true that there are too many errors in current work of this kind. One might suppose that astronomers, with whom the subject is a century old, would be open in small degree to this criticism. But it appears that some are content to learn how to use the method without showing any great curiosity regarding the underlying principles. It may not be without value to point out some common inaccuracies and errors that the writer has noticed.

In a recent number of the *Astronomische Nachrichten* we read the following definition: "The probable error means that where no force is at work the result will probably differ from zero by just this amount." Evidently this is a recurrence of the time-honored confusion between probable error and "most probable error." The former is such that the chance for the real error exceeding it is just equal to the chance of the real error being less than it; or, otherwise stated, if p be the probable error of an observed quantity, m, then there is an even chance that the true value of m lies between

$$m - p \qquad \text{and} \qquad m + p$$

The "most probable" error is always zero, for this is only another way of stating a fundamental assumption, that large errors are less likely to occur than smaller ones.

It should be noticed that there is another definition of probable error,—namely, that implied in the expression

$$\pm 0.6745 \sqrt{\frac{[v\,v]}{n\,(n-1)}}$$

which is not quite identical with the former. The two approach equality only when n, the number of observations, becomes

large. For example, if $n = 2$ the formula gives as the probable error of the mean about *one third* of $(m_2 - m_1)$, m_2 and m_1 being the two observations. It is easy to see, however, that the former definition leads to *one half* of $(m_2 - m_1)$ as the probable error. For we may have four cases, all equally likely to occur:—

(1) m_1 and m_2, each greater than the true value.
(2) m_1 greater, and m_2 less than the true value.
(3) m_2 greater, and m_1 less than the true value.
(4) m_1 and m_2, each less than the true value.

Consequently, there is just an even chance that the true value lies between

$$\frac{m_2 + m_1}{2} - \frac{m_2 - m_1}{2} \text{ and } \frac{m_2 + m_1}{2} + \frac{m_2 - m_1}{2}$$

According to the *former definition*, then, $\frac{m_2 - m_1}{2}$ must be the probable error of the mean.*

A misleading practice that one sees quite often is exemplified by the following. Suppose we have three determinations of a quantity with their probable errors:—

$$100 \pm 1$$
$$110 \pm 3$$
$$120 \pm 2$$

which we desire to combine into a final value. It is quite usual for the computer to obtain the weights from the indicated probable errors and to combine the three determinations accordingly into

$$104.5 \pm 0.86$$

This process ignores the existence of systematic errors (which in this case are so evidently present), and the result is open to criticism on two accounts: first, the adopted mean is not the best that can be gotten from the data; and, second, the probable error given is misleadingly small. Far more busi-

* A rough but very convenient way of estimating the probable error of the mean of a small number of observations (not over five or six) is to divide twice the range by three times the number of observations. By the *range* we mean the difference between the greatest and the least of the observed quantities. This will usually give a sufficiently close approximation to the expression $0.6745\sqrt{\frac{[v\,v]}{n\,(n-1)}}$.

nesslike would it be to take the simple mean of the three determinations and to estimate the probable error at 4 or 5.

Another irritating practice is the loose employment of the word "uncertainty." Thus, one astronomer, after computing a probable error of 0".012, states that the quantity to which it is attached "is subject to an uncertainty of perhaps 0".02." It is doubtful whether this writer could say what he means by "uncertainty," and surely the word will convey nothing definite to his readers. This usage should be abandoned, at least until a precise definition is given to the word; and even then it would be inadvisable to use it, for we already have two errors in common use, the *mean* and the *probable,* to indicate the reliability of a quantity. It would be convenient if one of these two could be employed to the exclusion of the other, for now it is necessary to specify which is being used in each particular case. German astronomers lean toward the mean error, and Americans toward the probable. The former obviates the necessity of multiplying by 0.6745; but this small advantage is outweighed, in our opinion, by the fact that the probable error has a concrete definition. Any person of fair intelligence can understand what probable errors are, even though he knows little of mathematics. The mean error, on the other hand, has an exceedingly technical definition. However, if one of these errors could be universally adopted, it would be a matter of small importance which should prevail.

Probably owing to the example set by GAUSS, it was formerly usual to give probable and mean errors to four or five significant figures. This practice has greatly decreased, but it has not entirely disappeared. In many cases two figures are ample, and it is difficult to imagine a case where more than three could be defended.

An error more serious than any of the above came under the writer's notice some time ago, and since then he has seen several other examples of it. The principle involved does not seem to be referred to in any of the standard works on the subject. Suppose we have a number of observation equations (equations of condition) involving several unknowns. Let these be solved by the method of least squares, and the resulting values of the unknowns be substituted in the original observation equations, giving the residuals in the usual way. Now,

considering each of the residuals in turn, it may be small for either of two reasons: first, because the corresponding observation was an accurate one; or, second, because of the way in which the coefficients of the unknown quantities enter. For example, if one of the unknowns rests for its determination principally upon only one of the observation-equations, then the residual for that equation will inevitably be small. The error to which the writer wishes to call attention results from the neglect of the possibility of a residual being small for any other reason than that the observation was an accurate one. A numerical example will make this clearer. We will suppose that we desire to find the declination (δ_0) at 1830 and the proper motion (μ) of a star which has been observed in 1800, 1830, 1840, and 1850. Assuming as a first approximation that all four observations are equally good, we solve by the method of least squares and compute the residuals:—

$$\begin{aligned} v_1 &= \delta_0 - 30\,\mu - \delta_1 \\ v_2 &= \delta_0 \qquad\quad\; - \delta_2 \\ v_3 &= \delta_0 + 10\,\mu - \delta_3 \\ v_4 &= \delta_0 + 20\,\mu - \delta_4 \end{aligned}$$

Suppose similar computations have been made for many other stars observed at the same dates, and that it is found that the average residual for the observations made in 1800 is about equal to the average residual for any of the three other dates. It would be a gross error to conclude from this that the observations made in 1800 are equally accurate with the others (yet just this has been done by some astronomers), because the proper motions depend so largely on the earliest observation that, other things being equal, we are bound to get small residuals for that date. It is not the writer's intention to go into the details of this point in this place. It will suffice to say that it is possible to compute the ratio between the actual errors of observation and the residuals, for any given set of equations. In the example above, these are respectively for the four equations, 0.11, 0.75, 0.68, and 0.46. If the residuals came out on the average proportionate to these ratios, then only could we conclude that the observations were of equal weight.

One other kind of error deserves mention before closing this note. In writing down our equations of condition we should be careful to let the measured quantities appear just

as we observe them. Thus, suppose we desire to find x and y from equations of the form

$$x + n_1 y = m_1$$
$$x + n_2 y = m_2$$
$$\text{etc., etc.}$$

These equations are evidently equivalent to

$$y + a_1 x = b_1$$
$$y + a_2 x = b_2$$
$$\text{etc., etc.,}$$

where $n_1 = \frac{1}{a_1}$, $n_2 = \frac{1}{a_2}$, etc.; and $m_1 = \frac{b_1}{a_1}$, $m_2 = \frac{b_2}{a_2}$, etc. In general, these two sets of equations, if solved by the method of least squares, will give two different values for x and two for y. The first set should be used if m_1, m_2 were observed, and the second set if b_1, b_2 were observed. Other examples can be constructed in which the fallacy is not quite so apparent as in the above. For an actual case the reader is referred to *Nature* (Vol. 63, page 489). FRANK SCHLESINGER.

Example of Personal Scale.—In connection with the note which appeared in the last number of these *Publications,* Mr. J. A. PARKHURST has kindly communicated the results of two counts: (1) of 1022 readings of a photometer scale; and (2) of 1148 readings of a micrometer. The first column below gives the estimated tenth; the other columns are percentages of the whole:—

	First Count. (Per cent.)	Second Count. (Per cent.)	Mean. (Per cent.)
0	17.5	14.0	15.7
1	7.1	10.3	8.7
2	12.0	11.9	12.0
3	9.7	9.4	9.6
4	6.2	5.3	5.8
5	7.4	8.1	7.8
6	4.5	6.4	5.4
7	11.3	9.8	10.5
8	16.9	13.0	14.9
9	7.4	11.8	9.6

This observer's scale is well pronounced and very symmetrical, the only exception in the latter respect being that 8's are preferred to 2's. F. S.

During 1903 four doctorates in astronomy were granted by universities in this country. This is equal to the average for the five preceding years. The recipients of these four degrees and the titles of their theses are: W. A. HAMILTON (University of Chicago), *On the Convergency of the Series in the Determination of the Elements of Parabolic Orbits and the Errors Introduced in the Elements by Imperfections of the Observations;* H. K. PALMER (University of California), *An Application of the Crossley Reflector of the Lick Observatory to the Study of Very Faint Spectra;* JOEL STEBBINS (University of California), *The Spectrum of o Ceti;* J. P. MCCALLIE (University of Virginia), *A Problem in Periodic Orbits, Second Order Perturbations of Jupiter and Saturn, Independent of the Eccentricities and of the Inclinations.*

From various sources, principally *Science,* we gather the following items: Dr. BURT L. NEWKIRK, a graduate ('97) of the University of Minnesota, has been appointed a computer on the orbits of the Watson asteroids at the University of California. Doctor NEWKIRK obtained his Ph.D. last year at Munich, his thesis being a photographic determination of the parallax of the central star in the annular nebula in *Lyra.*——Dr. CHARLES S. HOWE, Professor of Astronomy at the Case School of Applied Science, has been appointed president of that institution.——Professor E. C. PICKERING and Sir WILLIAM HUGGINS are among those who received honorary doctorates at the recent celebration at the University of Heidelberg.——Mr. ARTHUR R. HINKS, well known for his work in astronomical photography, has been promoted to the chief observership at Cambridge Observatory, England, in the room of Mr. GRAHAM, who has retired at an advanced age. Mr. W. E. HARTLEY has been appointed to the post made vacant by Mr. HINK's promotion.——M. LEBOEUF has been appointed Professor of Astronomy at the University of Besânçon.——Professors H. H. TURNER (of Oxford) and W. KAPTEYN (of Utrecht) have accepted invitations to visit this country as representatives of astronomy at the congress which is to be held at the St. Louis Fair. Among the mathematicians who have accepted similar invitations is POINCARÉ.——H. L. SMITH, Professor of Astronomy and Physics in Hobart College, died on August 1st.

PROSPER HENRY, astronomer *adjoint* at the Paris Observatory, died on July 25th, at Pomogen in the French Alps, as a result of exposure to extreme cold. The two brothers, PROSPER and PAUL, have been connected with the Paris Observatory for nearly forty years, and have taken a great part in the recent work of that institution. It is said that there is not an instrument in the observatory that the HENRYS have not made or remodeled. Very rarely has one of the brothers published anything apart from the other, and their work is so interwoven that probably no one knows just what is due to the one and what to the other. An illustration of this will be found in their discoveries of minor planets. The first of these is credited to PROSPER, the next to PAUL, the next to PROSPER, and so on through the whole fourteen. The HENRYS will no doubt be remembered longest for the share they took in bringing about the Astrophotographic Catalogue. Their other important work consists in the making of lenses and plane mirrors, asteroid discoveries, discovery of nebulæ in the *Pleiades,* method for reducing photographic star positions, and many other researches in astronomical photography.

F. S.

On May 2, 1903, ground was broken for the erection of a new observatory at Amherst College. The plans contemplate a structure one hundred and fifty feet long, east and west, surmounted by three domes, the central one to contain an equatorial refractor with a lens eighteen inches in diameter. This telescope is now in the process of construction by the Alvan Clark & Sons Corporation.

Number III of volume XLVIII of the Annals of Harvard College Observatory, containing a Provisional Catalogue of Variable Stars, forms a very welcome addition to variable-star literature. The last preceding catalogue, CHANDLER'S third, was published in 1896, and, as the number of variable stars has increased largely since that time, it had become impossible, unless one kept up a card catalogue, to keep properly informed of the advances in this work.

SCHÖNFELD'S first catalogue (1865) contains 113 stars;

CHANDLER'S third catalogue (1896), 393 stars; while in the present catalogue 1,227 stars are included, 509 of which, however, are the variables discovered by Professor BAILEY in clusters. In constructing this catalogue Professor PICKERING has departed from the guiding precepts used for other catalogues, and has included many stars whose variability has not yet been confirmed, claiming that it is better to include some stars not variable (definite names having not yet been assigned) than to exclude some which are variable. This position is certainly correct, especially for a provisional catalogue. One of the chief uses of a variable-star catalogue is to guide the observer in making up the observing list. It should therefore contain, in my opinion, not only a list of the certainly recognized variables, but a list, preferably, supplementary, of all stars ever suspected of variability, even though observations may have shown that for a considerable length of time the star was constant in light.

A new form of designation has been introduced into the catalogue, consisting of six figures which give the approximate position of the star. The first two figures are those of the hour, and the next two those of the minutes of Right Ascension of the star for 1900. The last two figures give the degree of Declination. South Declination is indicated by printing the figures in italics. The designation of the first star of the catalogue is *000339*, which means R. A. $0^h\ 3^m$, Dec. $-39°$. In defense of this system Professor PICKERING states as follows: "A new form of notation is proposed with much hesitation, but it has been found so convenient in actual use here that it has been retained. It aims to furnish a distinctive designation for any star, which will not be too long to be readily retained in the memory, and yet will give the position with sufficient precision to give the approximate setting for a visual or photographic telescope, or to decide whether the object is conveniently situated for observation."

It might be stated, however: *First.* It is not at all certain that six figures can be readily retained in the memory for any considerable length of time,—for instance, while one is making observations of the variable. I believe it will be generally admitted that the average person of astronomical training will retain four figures without effort. But when the number of

figures reaches six, I believe that, in general, a distinct mental effort will be necessary to retain them. Any one accustomed to using logarithms knows how much more mental effort is required in using six-place tables than in using four-place tables, although of course considerable of this added effort comes in making the interpolations.

Second. If the observer is working with a telescope larger than four inches, he will desire for setting purposes a declination more accurate than the nearest degree.

Third. It will be necessary for the observer in recording observations to adopt some symbol to denote South Declination. This will require added effort, and the omission of the symbol may lead to confusion between north and south stars. It should be stated, however, that this last is a theoretical rather than a practical objection, for in the catalogue under consideration there are only two instances of two stars, one north and one south, having the same designation.

Fourth. It will very seldom happen that one will have at hand the designation of the variable and not its coordinates also; and if one wishes to see whether or not the star is in position for observation, it will be more convenient to refer to the columns of Right Ascension and Declination than to the designation-number.

Fifth. Although good for a great many years, the position given in the designation-number, because referred to the epoch of 1900, would ultimately not be sufficiently exact for setting purposes. Our successors of the twenty-first and following centuries may get tired of referring to the epoch of 1900.

CHANDLER'S system of designating variable stars has been pretty generally adopted, and should be retained, I believe, until one is proposed which is distinctly and unmistakably better.

The new catalogue contains two columns not found in CHANDLER'S catalogues. The one gives the class to which the variable belongs, the classification being that proposed by Professor PICKERING in 1880; the other designates the spectrum of the variable, the notation employed being the same as that used in the DRAPER catalogue.

Copious notes and several supplementary tables follow the catalogue.

An examination of this catalogue reveals the vast amount of work that variable-star observers have before them. Some of the columns are conspicuous for their incompleteness. There are listed in the catalogue, excluding new stars and variables in clusters, 286 stars whose periods are only approximately known or wholly unknown.

S. D. T.

PUBLICATIONEN DES ASTROPHYSIKALISCHEN OBSERVATORIUMS KÖNIGSTUHL-HEIDELBERG, HERAUSGEGEBEN VON DR. MAX WOLF. ERSTER BAND. KARLSRUHE, G. BRAUN. 1902.

This interesting volume is the first of the publications of this observatory under its new régime. Wisely limiting itself to one chosen field, the observatory has produced a volume which entitles it to a high rank as an institution of definitely concentrated research. A full description of the buildings and instruments is left for a subsequent volume. Provided as it is with two Voigtländer 6-inch cameras and the two great 16-inch Bruce photographic lenses, the Königstuhl Observatory is singularly well equipped for stellar and planetary photography. It is unnecessary here to mention Dr. WOLF's successful work on the asteroids.

As a field in which comparatively little has been done, the Königstuhl Observatory has fixed upon the photography and accurate cataloguing of nebulæ as its main work, and it is to this subject that the present volume is largely devoted. A rapid and accurate method was needed to measure the positions of the large number of small nebulæ discovered photographically, and a parallactic measuring-engine on the lines proposed by KAPTEYN was procured. In connection with the First Königstuhl List of 154 Nebulæ, a description of the instrument is given by Dr. WOLF, with an investigation of the errors of the screws and circles. The method of the parallactic measuring-engine is, in brief, as follows: Let us suppose the measuring-microscope to be mounted as a small equatorial and the photographic plate to be placed at a distance from it equal to the focal length of the lens with which it was taken. Then, if the photographic plate is placed in a position which, as seen

from the little measuring-equatorial, is of the same hour-angle and declination as was the region of the sky at the time of the exposure, we can with the equatorially-mounted microscope read off the positions of objects directly in differences of Right Ascension and Declination, much as would be done with an equatorially-mounted telescope fitted with a filar micrometer. In practice the equatorially-mounted microscope is placed with its polar axis horizontal, and is provided with a third vertical axis for convenience.

Dr. SCHWASSMAN, after treating very fully the theory of the apparatus, has used it in the preparation of the Second Königstuhl List of 301 Nebulæ, employing the full and rigorous methods demanded by the theory in the case of plates of large area. This second list contains the positions for 1900 of 301 nebulæ in *Virgo*. The employment of the full formulæ must have consumed a very great deal of time in both measurement and reduction, and such refinement would seem entirely unnecessary in objects of the definition of an ordinary nebular patch, and in plates of this area, where the images of ninth-magnitude stars are about 47″ in diameter. Dr. SCHWASSMAN finds for the mean error of a position

$$m_\alpha = \pm 0^s.16 \qquad m_\delta = \pm 1''.8$$

and a comparison with micrometer results shows that the accuracy of the coordinates in Right Ascension is of the same order of magnitude as that of direct micrometer measures, while the accuracy in Declination is slightly inferior.

In a later paper in the same volume Dr. WOLF gives the Third Königstuhl List of Nebulæ, comprising 1,528 nebulæ near the pole of the Milky Way. By using more approximate methods, and neglecting the full and rigorous formulæ of KAPTEYN'S method, a great saving of time was effected without appreciable loss of accuracy. Dr. WOLF estimates the error of a position in the third list as somewhat larger than $0^s.1$ and $1''.0$. As to the time consumed, about eighty-two hours were spent in the actual measurement, so that the positions of 1,528 nebulæ to the above degree of accuracy were secured in a few weeks, whereas had the full and rigorous method employed by Dr. SCHWASSMAN been used it would have been several years. Of these 1,528 nebulæ only one in nineteen, or about five per cent,

were previously known, confirming KEELER'S conclusions (*Astrophysical Journal,* vol. XI, p. 325) as to the enormous number of these small objects. A map is given showing the density of the grouping of these nebulæ about the pole of the Milky Way, and Dr. WOLF states that he knows of no other region where nebulæ are so thickly clustered, there being at the densest portion no fewer than seventy in one sixteenth of a square degree. Another point of interest brought out in this investigation is, that the large majority of the position-angles of the axes of "long" nebulæ are grouped about 60°.

Dr. CARNERA investigates the light-curve of the variable star *S Leonis* by photographic methods. The method used was that of the measurement of the diameters of the star-disks on the photographic plates. Numerous precautions were taken to avoid the errors which may so easily arise in this method, and plates were taken of the region of the variable, the *Pleiades,* and the Pole at the same angle of elevation above the horizon. CHANDLER'S period and light-curve is confirmed. Dr. CARNERA gives also a number of photographic positions of *Eros.*

Mr. A. KOPFF, in an interesting paper, gives an investigation from the negatives of two regions of the phenomenon, first noted by HERSCHEL, that the regions about many nebulæ are relatively starless. The regions investigated are those about the Nebula in *Orion* and the great "America" Nebula. A count of the stars was made, showing that each of these is surrounded by a veritable star-desert, the *Orion* Nebula lying near the northwest and the America Nebula near the northeast end of the relatively starless tracts surrounding them.

H. D. C.